

(19)



Europäisches Patentamt

European Patent Office

Office européen des brevets



(11) EP 1 239 458 A2

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
11.09.2002 Bulletin 2002/37

(51) Int Cl.7: G10L 15/02

(21) Application number: 02005252.8

(22) Date of filing: 08.03.2002

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE TR  
Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 08.03.2001 JP 2001064286

(71) Applicant: NEC CORPORATION  
Tokyo (JP)

(72) Inventor: Iso, Kenichi, NEC Corporation  
Minato-ku, Tokyo (JP)

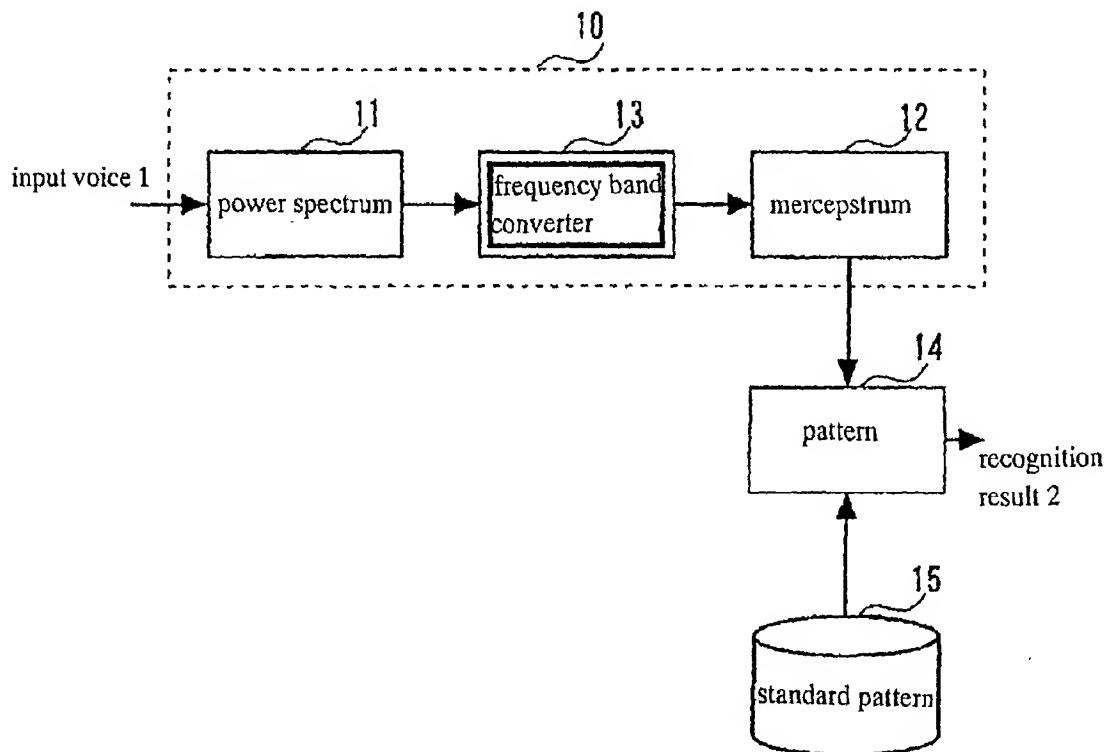
(74) Representative: VOSSIUS & PARTNER  
Siebertstrasse 4  
81675 München (DE)

### (54) Voice recognition system, standard pattern preparation system and corresponding methods

(57) A voice recognition system for recognizing an input voice of a narrow frequency band. The voice recognition system includes : a frequency band converting

unit for converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

FIG. 3



## Description

**[0001]** The present invention relates to a voice recognition system and a standard pattern preparation system for preparing standard patterns to be used for the voice recognition process by the voice recognition system as well as a method of preparing the standard patterns and a computer program for preparing the standard patterns, and more particularly to a voice recognition system for recognizing a narrow-band frequency voice such as a telephone voice recognition.

**[0002]** FIG. 1 is a block diagram illustrative of the conventional voice recognition system. The conventional voice recognition system includes a characteristic extraction unit 100 and a pattern reference unit 103. The characteristic extraction unit 100 receives an input of a voice 105 and converts the voice into a characteristic vector time series. The pattern reference unit 103 receives the characteristic vector time series and compares the characteristic vector time series with a standard pattern 104 for the voice recognition before the pattern reference unit 103 outputs a voice recognition result 106. This conventional voice recognition system is addressed in 1995, entitled "the fundamentals of the voice recognition" NTT Advanced Technology.

**[0003]** A mercepstrum characteristic extraction may be available for the characteristic extraction unit 100. The characteristic extraction unit 100 further includes a power spectrum calculation unit 101 for calculating a power spectrum in a short term of the input voice 105 and a mercepstrum calculation unit 102 receiving the power spectrum from the power spectrum calculation unit 101 and performing a mer conversion and a cosine conversion of a logarithm of the power spectrum, thereby extracting a mercepstrum characteristic quantity.

**[0004]** FIG. 2 is a block diagram illustrative of a conventional standard pattern preparation system. The conventional standard pattern preparation system prepares the above described standard pattern 104 to be referred by the above described conventional voice recognition system shown in FIG. 1. The conventional standard pattern preparation system includes a characteristic extraction unit 200 and a standard pattern preparation unit 204. The characteristic extraction unit 200 further includes a power spectrum calculation unit 201 for calculating a power spectrum in a short term of a learning voice signal from a learning voice storing unit 203, and a mercepstrum calculation unit 202 receiving the power spectrum from the power spectrum calculation unit 201 and performing a mer conversion and a cosine conversion of a logarithm of the power spectrum, thereby extracting a mercepstrum characteristic quantity.

**[0005]** The standard pattern preparation unit 204 receives the mercepstrum characteristic quantity from the mercepstrum calculation unit 202 and prepares a standard pattern. The standard pattern is stored in a standard pattern storing unit 205.

**[0006]** With reference again to FIG. 1 the process for recognition of the narrow band frequency voice such as the telephone voice by the conventional voice recognition system will be described.

**[0007]** The telephone voice has a narrow frequency band, and is likely to receive a substantive influence by noises, for which reason it is generally difficult to recognize the voice. The frequency band of the telephone voice is ranged from 300 Hz to 3400 Hz. A first formant of the vowel or the primary characteristic frequency region is important for the voice recognition. This first formant of the vowel or the primary characteristic frequency region exists under 300 Hz, depending on a speaker. In this case, the voice signal entered from the telephone terminal may be free of the first formant of the vowel or the primary characteristic frequency region under 300 Hz.

**[0008]** A frequency range of the friction noise may often be over 3000 Hz. In this case, the voice signal entered from the telephone terminal may be free of the friction noise.

**[0009]** The restriction on the frequency band causes that the recognition of the telephone voice with the narrow frequency band is lower in accuracy than the recognition of the microphone voice with a wide frequency band.

**[0010]** JP-A- 2000-250577 discloses the following conventional technique for improving the frequency characteristic of the voice with the narrow frequency band entered from the microphone. This conventional technique prevents any lack of the voice information and also improves the voice recognition characteristic in the presence of the noises. A characteristic vector is selected by a first code book from a voice input pattern as received by a second voice receiver. A correction vector is selected from a second code book in correspondence with the index of the selected vector. Both the characteristic vector and the correction vector are then added to presume the characteristic vector of the voice received by the second receiver which ensures a higher voice-receiving sensitivity in a wide frequency band than the first receiver.

**[0011]** In the above circumstances, the development of a novel voice recognition system is desirable.

**[0012]** Accordingly, it is an object of the present invention to provide a novel recognition system free from the above problems.

**[0013]** It is a further object of the present invention to provide a novel recognition system exhibiting such a high performance of voice recognition in a narrow frequency band as closely to the performance in the narrow frequency band.

**[0014]** It is a still further object of the present invention to provide a novel standard pattern preparation system for preparing standard patterns to be used for the voice recognition process by the voice recognition system free from the above problems.

**[0015]** It is yet a further object of the present invention

to provide a novel standard pattern preparation system for preparing standard patterns to be used for the voice recognition process by the voice recognition system exhibiting such a high performance of voice recognition in a narrow frequency band as closely to the performance in the narrow frequency band.

**[0016]** It is yet a further object of the present invention to provide a method of preparing the standard patterns free from the above problems.

**[0017]** It is yet a further object of the present invention to provide a method of preparing the standard patterns exhibiting such a high performance of voice recognition in a narrow frequency band as closely to the performance in the narrow frequency band.

**[0018]** It is yet a further object of the present invention to provide a computer program for preparing the standard patterns free from the above problems.

**[0019]** It is yet a further object of the present invention to provide a computer program for preparing the standard patterns exhibiting such a high performance of voice recognition in a narrow frequency band as closely to the performance in the narrow frequency band.

**[0020]** The present invention provides a voice recognition system for recognizing an input voice of a narrow frequency band. The voice recognition system includes : a frequency band converting unit for converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

**[0021]** A first aspect of the present invention is a voice recognition system for recognizing an input voice of a narrow frequency band. The voice recognition system includes : a frequency band converting unit for converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

**[0022]** It is preferable that the voice recognition system further includes : a power spectrum calculating unit for calculating power spectrums of the input voice of the narrow frequency band, and that the frequency band converting unit further comprises : a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice ; an expansion coefficient calculating unit for calculating expansion coefficients that the power spectrums calculated by the power spectrum calculating unit are expanded by the specific vectors ; and a frequency band expansion unit for calculating additional power spectrums in a lack frequency band by use of the expansion coefficients calculated by the expansion coefficient calculating unit, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, and the frequency band expansion unit combining the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band calculated by the power spectrum calculating unit to prepare

pseudo power spectrums of the pseudo voice of the wide frequency band.

**[0023]** It is also preferable to further include : a pattern reference unit for receiving an output from the frequency band converting unit and comparing the output with a standard pattern for carrying out the voice recognition.

**[0024]** It is also preferable to further include : a standard pattern storing unit for storing the standard patterns for allowing the pattern reference unit to compare the output with the standard pattern.

**[0025]** It is also preferable to further include : a standard pattern preparation unit for receiving an output from the frequency band converting unit and preparing a standard pattern based on the output for carrying out the voice recognition.

**[0026]** It is also preferable to further include : a standard pattern storing unit for storing the standard pattern prepared by the standard pattern preparation unit.

**[0027]** It is also preferable that the frequency band converting unit further comprises : a mercepstrum calculating unit for calculating a mercepstrum characteristic quantity based on the pseudo power spectrum prepared by the frequency band expansion unit.

**[0028]** A second aspect of the present invention is a standard pattern preparation system for preparing a standard pattern to recognize an input voice of a narrow frequency band. The standard pattern preparation system includes : a frequency band converting unit for converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

**[0029]** It is also preferable that the standard pattern preparation system further includes : a power spectrum calculating unit for calculating power spectrums of the input voice of the narrow frequency band, and that the frequency band converting unit further comprises : a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice ; an expansion coefficient calculating unit for calculating expansion coefficients that the power spectrums calculated by the power spectrum calculating unit are expanded by the specific vectors ; and a frequency band expansion unit for calculating additional power spectrums in a lack frequency band by use of the expansion coefficients calculated by the expansion coefficient calculating unit, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, and the frequency band expansion unit

combining the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band calculated by the power spectrum calculating unit to prepare pseudo power spectrums of the pseudo voice of the wide frequency band.

**[0030]** It is also preferable to further include : a pattern reference unit for receiving an output from the frequency band converting unit and comparing the output with a standard pattern for carrying out the voice recognition.

**[0031]** It is also preferable to further include : a standard pattern storing unit for storing the standard patterns for allowing the pattern reference unit to compare the output with the standard pattern.

**[0032]** It is also preferable to further include : a standard pattern preparation unit for receiving an output from the frequency band converting unit and preparing a standard pattern based on the output for carrying out the voice recognition.

**[0033]** It is also preferable to further include : a standard pattern storing unit for storing the standard pattern prepared by the standard pattern preparation unit.

**[0034]** It is also preferable that the frequency band converting unit further comprises : a mercepstrum calculating unit for calculating a mercepstrum characteristic quantity based on the pseudo power spectrum prepared by the frequency band expansion unit.

**[0035]** A third aspect of the present invention is a characteristic extraction system for extracting a mercepstrum characteristic quantity of an input voice of a narrow frequency band. The characteristic extraction system includes : a power spectrum calculating unit for calculating power spectrums of the input voice of the narrow frequency band ; a frequency band converting unit for converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band ; and a mercepstrum calculating unit for calculating a mercepstrum characteristic quantity based on the pseudo voice of the wide frequency band.

**[0036]** It is also preferable that the frequency band converting unit further comprises : a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice ; an expansion coefficient calculating unit for calculating expansion coefficients that the power spectrums calculated by the power spectrum calculating unit are expanded by the specific vectors ; and a frequency band expansion unit for calculating additional power spectrums in a lack frequency band by use of the expansion coefficients calculated by the expansion coefficient calculating unit, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, and the frequency band expansion unit combining the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band calculated by the power spectrum calculating unit to prepare pseudo power spectrums of the pseudo voice of the wide frequency band.

**[0037]** A fourth aspect of the present invention is a voice recognition method for recognizing an input voice of a narrow frequency band. The voice recognition method includes the steps of: converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

**[0038]** It is also preferable to further include : calculating power spectrums of the input voice of the narrow frequency band ; calculating expansion coefficients that the power spectrums are expanded by specific vectors ; and calculating additional power spectrums in a lack frequency band by use of the expansion coefficients, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, so as to combine the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band, thereby to prepare pseudo power spectrums of the pseudo voice of the wide frequency band.

**[0039]** It is also preferable to further include : calculating a mercepstrum characteristic quantity based on the pseudo power spectrum.

**[0040]** It is also preferable to further include : comparing the mercepstrum characteristic quantity with a standard pattern for carrying out the voice recognition.

**[0041]** A fifth aspect of the present invention is a standard pattern preparation method for preparing a standard pattern to recognize an input voice of a narrow frequency band. The standard pattern preparation method includes the steps of : converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band.

**[0042]** It is also preferable to further include : calculating power spectrums of the input voice of the narrow frequency band ; calculating expansion coefficients that the power spectrums are expanded by specific vectors ; and calculating additional power spectrums in a lack frequency band by use of the expansion coefficients, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, and combining the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band to prepare pseudo power spectrums of the pseudo voice of the wide frequency band.

**[0043]** It is also preferable to further include : calculating a mercepstrum characteristic quantity based on the pseudo power spectrum.

**[0044]** It is also preferable to further include : comparing the mercepstrum characteristic quantity with a standard pattern for carrying out the voice recognition.

**[0045]** A sixth aspect of the present invention is a characteristic extraction method for extracting a mercepstrum characteristic quantity of an input voice of a narrow frequency band. The characteristic extraction method includes the steps of : calculating power spectrums of the input voice of the narrow frequency band ; converting the input voice of the narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of the narrow frequency band and which is wider than the narrow frequency band ; and calculating a mercepstrum characteristic quantity based on the

pseudo voice of the wide frequency band.

[0046] It is also preferable to further include : calculating expansion coefficients that the power spectrums are expanded by specific vectors ; and calculating additional power spectrums in a lack frequency band by use of the expansion coefficients, where the lack frequency band is covered by the wide frequency band but not covered by the narrow frequency band, and combining the additional power spectrums in the lack frequency band into the power spectrum of the narrow frequency band to prepare pseudo power spectrums of the pseudo voice of the wide frequency band.

[0047] In accordance with the present invention, the characteristic extraction unit for analyzing the input voice signal and extracting the characteristic quantity further includes a frequency band conversion unit for converting the narrow frequency band power spectrum into a wide frequency band power spectrum.

[0048] The frequency band conversion is executed for preparing the standard pattern in order to reduce the mismatch between the standard pattern and the input voice.

[0049] The characteristic extraction unit for analyzing the input voice signal and extracting the characteristic quantity further includes a frequency band conversion unit for converting the narrow frequency band power spectrum into a wide frequency band power spectrum. The frequency band conversion unit further includes a characteristic vector storing unit for storing characteristic vectors of a power spectrum of the wide frequency band voice. A plurality of power spectrums of the wide frequency band voice are gathered for carrying out a main component analysis to the power spectrums in order to prepare the characteristic vectors. The characteristic vectors are then stored in the characteristic vector storing unit. Any power spectrums of the wide frequency band voice are represented by a linear combination of the characteristic vectors.

[0050] Short time power spectrums of a narrow frequency band voice of the input voice signal are calculated by a power spectrum calculating unit. The frequency band conversion unit furthermore includes an expansion coefficient calculation unit for calculating an expansion coefficient, wherein the calculated power spectrum of the narrow frequency band is expanded by the characteristic vector of the wide frequency band stored in the characteristic vector storing unit.

[0051] The frequency band conversion unit moreover includes a frequency band expansion unit for calculating a pseudo wide frequency band voice power spectrum which is expanded by the expansion coefficient from the input narrow frequency band power spectrum. In details, the frequency band expansion unit calculates a power spectrum of a lack frequency band by utilizing the expansion coefficient, and combines the calculated power spectrum of the lack frequency band into the already calculated narrow frequency band spectrum, so that the frequency band expansion unit calculates the pseudo

wide frequency band voice power spectrum.

[0052] The voice recognition system further includes a mercepstrum calculating unit for calculating the mercepstrum as the mercepstrum characteristic quantity by utilizing the above described pseudo wide frequency band voice power spectrum.

[0053] The expansion coefficient calculation process by the expansion coefficient calculation unit and the frequency band expansion process by the frequency band expansion unit may be realized by programs executed by the computer. The programs may be stored in a computer-readable storage medium such as a semiconductor memory or a magnetic recording medium.

[0054] The above and other objects, features and advantages of the present invention will be apparent from the following descriptions.

[0055] Preferred embodiments according to the present invention will be described in detail with reference to the accompanying drawings.

[0056] FIG. 1 is a block diagram illustrative of the conventional voice recognition system.

[0057] FIG. 2 is a block diagram illustrative of a conventional standard pattern preparation system.

[0058] FIG. 3 is a block diagram illustrative of a first novel voice recognition system in a first embodiment in accordance with the present invention.

[0059] FIG. 4 is a block diagram illustrative of the structure of the characteristic extraction unit included in the novel voice recognition system of FIG. 3.

[0060] FIG. 5 is a block diagram illustrative of a first novel voice recognition system in a second embodiment in accordance with the present invention.

[0061] FIG. 6 is a block diagram illustrative of the structure of the characteristic extraction unit included in the novel voice recognition system of FIG. 5.

#### FIRST EMBODIMENT:

[0062] A first embodiment according to the present invention will be described in detail with reference to the drawings. FIG. 3 is a block diagram illustrative of a first novel voice recognition system in a first embodiment in accordance with the present invention.

[0063] The voice recognition system includes a characteristic extraction unit 10, a pattern reference unit 14. The characteristic extraction unit 10 receives an input of a voice 1 and performs the following processes. The pattern reference unit 2 receives the output from the characteristic extraction unit 10 and compares the output with a standard pattern 15 for the voice recognition.

[0064] The characteristic extraction unit 10 is provided for analyzing the input voice signal 1 and extracting the characteristic quantity. The characteristic extraction unit 10 further includes a frequency band conversion unit 13 for converting the narrow frequency band power spectrum into a wide frequency band power spectrum. FIG. 4 is a block diagram illustrative of the structure of the characteristic extraction unit included in the novel

voice recognition system of FIG. 3.

[0065] The frequency band conversion unit 13 further includes a characteristic vector storing unit 33 for storing characteristic vectors of a power spectrum of the wide frequency band voice. A plurality of power spectrums of the wide frequency band voice are gathered for carrying out a main component analysis to the power spectrums in order to prepare the characteristic vectors. The characteristic vectors are then stored in the characteristic vector storing unit 33. Any power spectrums of the wide frequency band voice are represented by a linear coupling of the characteristic vectors.

[0066] Short time power spectrums of a narrow frequency band voice of the input voice signal are calculated by a power spectrum calculating unit.

[0067] The frequency band conversion unit 13 furthermore includes an expansion coefficient calculation unit 31 for calculating an expansion coefficient, wherein the calculated power spectrum of the narrow frequency band is expanded by the characteristic vector of the wide frequency band stored in the characteristic vector storing unit 33.

[0068] The frequency band conversion unit 13 moreover includes a frequency band expansion unit 32 for calculating a pseudo wide frequency band voice power spectrum which is expanded by the expansion coefficient from the input narrow frequency band power spectrum. In details, the frequency band expansion unit 32 calculates a power spectrum of a lack frequency band by utilizing the expansion coefficient, and combines the calculated power spectrum of the lack frequency band into the already calculated narrow frequency band spectrum, so that the frequency band expansion unit 32 calculates the pseudo wide frequency band voice power spectrum.

[0069] With reference back to FIG. 3, the characteristic extraction unit 10 further includes a mercepstrum calculating unit 12 for calculating the mercepstrum as the mercepstrum characteristic quantity by utilizing the above described pseudo wide frequency band voice power spectrum.

[0070] The expansion coefficient calculation process by the expansion coefficient calculation unit 31 and the frequency band expansion process by the frequency band expansion unit 32 may be realized by programs executed by the computer. The programs may be stored in a computer-readable storage medium such as a semiconductor memory or a magnetic recording medium.

[0071] Operations of the above described novel voice recognition system will be described. The input voice 1 having a restricted narrow frequency band is inputted into the power spectrum calculating unit 11. The power spectrum calculating unit 11 calculates the narrow band power spectrums.

$$\{N(k)\}, k = K_1 - K_2 \quad (1)$$

where  $N(k)$  is the power spectrum of the voice at a frequency "k", and  $K_1 - K_2$  correspond to upper and lower limits of the narrow band frequencies. In case of the telephone voice,  $K_1$  is 300Hz and  $K_2$  is 3400Hz.

5 [0072] The frequency band conversion unit 13 converts the narrow frequency band power spectrum into the wide frequency band power spectrum.

$$10 \quad \{W(k)\}, k = K_0 - K_3 \quad (2)$$

where  $W(k)$  is the power spectrum of the voice at a frequency "k", and  $K_0 - K_3$  correspond to upper and lower limits of the wide band frequencies.

$$15 \quad K_0 < K_1 < K_2 < K_3 \quad (3)$$

[0073] The mercepstrum calculating unit 12 calculates the mercepstrum by utilizing the wide frequency band power spectrum  $W(k)$  without utilizing the narrow frequency band power spectrum  $N(k)$ . The voice informations in the narrow frequency band  $K_0-K_1$  and the wide frequency band  $K_2-K_3$  are utilized for improving the accuracy of the voice recognition.

[0074] With reference to FIG. 4, the characteristic vector storing unit 33 stores the characteristic vectors of the power spectrum of the wide frequency band voice.

$$30 \quad \{\phi(i, k)\}, i = 1 - M, k = K_0 - K_3 \quad (4)$$

where "M" is the number of the characteristic vectors which have already been stored in the characteristic vector storing unit 33. The characteristic vectors have been prepared by gathering power spectrums of the wide frequency voice and the main component analysis.

[0075] The wide frequency band power spectrum  $W(k)$  may be represented by a linear combination of the characteristic vectors.

$$W(k) = \sum_{i=1-M} c(i) \phi(i, k), k = K_0 - K_3 \quad (5)$$

45 where  $c(i)$  is the expansion coefficient.

[0076] The expansion coefficient calculation unit 31 calculates the expansion coefficient, wherein the narrow frequency band power spectrum  $\{N(k)\}$  from the power spectrum calculating unit 11 is expanded by the characteristic vector of the wide frequency band power spectrum stored in the characteristic vector storing unit 33.

$$50 \quad \{N(k)\}, k = K_1 - K_2 \quad (6)$$

[0077] The expansion coefficient  $b(i)$  is calculated to minimize the squared error "E" represented by:

$$E = \sum_{k=K1, K2; i=1-M} \{N(k) - \sum b(i)\}^2 \quad (7)$$

**[0078]** The frequency band expansion unit 32 receives the expansion coefficient  $b(i)$  from the expansion coefficient calculation unit 31, and calculates the wide frequency band power spectrum  $W(k)$  by use of the following equation including the expansion coefficient  $b(i)$ .

$$W(k) = \sum_{i=1-M} b(i)\phi(i,k), k = K0 - K3 \quad (8)$$

**[0079]** The mercepstrum calculating unit 12 receives the wide frequency band power spectrum  $W(k)$  from the frequency band expansion unit 32, and calculates the mercepstrum characteristic quantity.

**[0080]** In the above description, the expansion coefficient  $b(i)$  is calculated to minimize the above described squared error "E". Alternatively, the expansion coefficient  $b(i)$  may optionally be calculated to maximize the likelihood of Hidden Markov Model to the input voice.

**[0081]** It is also possible that both the narrow frequency band power spectrum and the wide frequency band power spectrum are calculated from the voice received in both the narrow and wide frequency bands. In the narrow frequency band, the narrow frequency band power spectrum is used. For the lack frequency region in the narrow frequency band, parts of the wide frequency band power spectrum corresponding to the lack frequency region are combined into the narrow frequency band power spectrum, so that the combination of the narrow frequency band power spectrum with the parts of the wide frequency band power spectrum form the wide frequency band power spectrum. The wide frequency band power spectrum is then subjected to the main component analysis to obtain the specific vectors. These specific vectors may be used for calculating the wide frequency band power spectrum  $W(k)$ .

#### SECOND EMBODIMENT:

**[0082]** A second embodiment according to the present invention will be described in detail with reference to the drawings. FIG. 5 is a block diagram illustrative of a first novel voice recognition system in a second embodiment in accordance with the present invention.

**[0083]** The voice recognition system includes a learning voice storing unit 24, a characteristic extraction unit 20, a standard pattern preparation unit 25. The characteristic extraction unit 10 receives an input of a voice 1 and performs the following processes. The standard pattern preparation unit 25 receives the output from the characteristic extraction unit 20 and prepares a standard pattern 26 for the voice recognition.

**[0084]** The characteristic extraction unit 20 further includes a frequency band conversion unit 23 for converting the narrow frequency band power spectrum into a

wide frequency band power spectrum. FIG. 6 is a block diagram illustrative of the structure of the characteristic extraction unit included in the novel voice recognition system of FIG. 5.

**[0085]** The frequency band conversion unit 23 further includes the characteristic vector storing unit 33 for storing characteristic vectors of a power spectrum of the wide frequency band voice. A plurality of power spectrums of the wide frequency band voice are gathered for carrying out a main component analysis to the power spectrums in order to prepare the characteristic vectors. The characteristic vectors are then stored in the characteristic vector storing unit 33. Any power spectrums of the wide frequency band voice are represented by a linear coupling of the characteristic vectors. Short time power spectrums of a narrow frequency band voice of the input voice signal are calculated by a power spectrum calculating unit.

**[0086]** The frequency band conversion unit 23 furthermore includes the expansion coefficient calculation unit 31 for calculating an expansion coefficient, wherein the calculated power spectrum of the narrow frequency band is expanded by the characteristic vector of the wide frequency band stored in the characteristic vector storing unit 33.

**[0087]** The frequency band conversion unit 23 moreover includes the frequency band expansion unit 32 for calculating a pseudo wide frequency band voice power spectrum which is expanded by the expansion coefficient from the input narrow frequency band power spectrum. In details, the frequency band expansion unit 32 calculates a power spectrum of a lack frequency band by utilizing the expansion coefficient, and combines the calculated power spectrum of the lack frequency band into the already calculated narrow frequency band spectrum, so that the frequency band expansion unit 32 calculates the pseudo wide frequency band voice power spectrum.

**[0088]** With reference back to FIG. 5, the characteristic extraction unit 20 further includes the mercepstrum calculating unit 22 for calculating the mercepstrum as the mercepstrum characteristic quantity by utilizing the above described pseudo wide frequency band voice power spectrum. The structure of the frequency band conversion unit 23 is substantially the same as in the first embodiment with reference to FIG. 4.

**[0089]** The standard pattern preparation unit 25 prepares the standard pattern based on the wide frequency band power spectrums of the voice to improve the accuracy of the standard pattern.

**[0090]** The learning voice storing unit 24 may store both the narrow frequency band voice and the wide frequency band voice. If the wide frequency band voice is stored in the learning voice storing unit 24, then a narrow frequency part is extracted from the wide frequency band power spectrum, and this extracted narrow frequency part is then converted by the frequency band conversion unit 23 into the pseudo wide frequency band

power spectrum.

[0091] Although the invention has been described above in connection with several preferred embodiments therefor, it will be appreciated that those embodiments have been provided solely for illustrating the invention, and not in a limiting sense. Numerous modifications and substitutions of equivalent materials and techniques will be readily apparent to those skilled in the art after reading the present application, and all such modifications and substitutions are expressly understood to fall within the true scope and spirit of the appended claims.

### Claims

1. A voice recognition system for recognizing an input voice of a narrow frequency band, said voice recognition system including:

a frequency band converting unit for converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band.

2. The voice recognition system as claimed in claim 1, characterized in that said voice recognition system further includes : a power spectrum calculating unit for calculating power spectrums of said input voice of said narrow frequency band, and  
characterized in that said frequency band converting unit further comprises:

a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice ;  
an expansion coefficient calculating unit for calculating expansion coefficients that said power spectrums calculated by said power spectrum calculating unit are expanded by said specific vectors ; and

a frequency band expansion unit for calculating additional power spectrums in a lack frequency band by use of said expansion coefficients calculated by said expansion coefficient calculating unit, where said lack frequency band is covered by said wide frequency band but not covered by said narrow frequency band, and said frequency band expansion unit combining said additional power spectrums in said lack frequency band into said power spectrum of said narrow frequency band calculated by said power spectrum calculating unit to prepare pseudo power spectrums of said pseudo voice of said wide frequency band.

3. A standard pattern preparation system for preparing a standard pattern to recognize an input voice of a narrow frequency band, said standard pattern preparation system including :

5 a frequency band converting unit for converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band.

4. The standard pattern preparation system as claimed in claim 3, characterized in that said standard pattern preparation system further includes : a power spectrum calculating unit for calculating power spectrums of said input voice of said narrow frequency band, and

10 characterized in that said frequency band converting unit further comprises:

a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice;

15 an expansion coefficient calculating unit for calculating expansion coefficients that said power spectrums calculated by said power spectrum calculating unit are expanded by said specific vectors ; and

20 a frequency band expansion unit for calculating additional power spectrums in a lack frequency band by use of said expansion coefficients calculated by said expansion coefficient calculating unit, where said lack frequency band is covered by said wide frequency band but not covered by said narrow frequency band, and said frequency band expansion unit combining said additional power spectrums in said lack frequency band into said power spectrum of said narrow frequency band calculated by said power spectrum calculating unit to prepare pseudo power spectrums of said pseudo voice of said wide frequency band.

- 30 45 5. The standard pattern preparation system as claimed in claim 2, 3 or 4 further including:

25 a pattern reference unit for receiving an output from said frequency band converting unit and comparing said output with a standard pattern for carrying out the voice recognition.

- 35 50 55 6. The standard pattern preparation system as claimed in claim 5, further including :

a standard pattern storing unit for storing said standard patterns for allowing said pattern reference unit to compare said output with said

standard pattern.

7. The standard pattern preparation system as claimed in claim 2, 3 or 4 further including:

a standard pattern preparation unit for receiving an output from said frequency band converting unit and preparing a standard pattern based on said output for carrying out the voice recognition.

8. The standard pattern preparation system as claimed in claim 7, further including :

a standard pattern storing unit for storing the standard pattern prepared by the standard pattern preparation unit.

9. The system as claimed in any one of claim 2 to 8, wherein said frequency band converting unit further comprises:

a mercepstrum calculating unit for calculating a mercepstrum characteristic quantity based on said pseudo power spectrum prepared by said frequency band expansion unit.

10. A characteristic extraction system for extracting a mercepstrum characteristic quantity of an input voice of a narrow frequency band, said characteristic extraction system including :

a power spectrum calculating unit for calculating power spectrums of said input voice of said narrow frequency band ;  
a frequency band converting unit for converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band ; and  
a mercepstrum calculating unit for calculating a mercepstrum characteristic quantity based on said pseudo voice of said wide frequency band.

11. The characteristic extraction system as claimed in claim 10, **characterized in that** said frequency band converting unit further comprises :

a specific vector storing unit for storing specific vectors of power spectrums of a wide frequency band voice ;  
an expansion coefficient calculating unit for calculating expansion coefficients that said power spectrums calculated by said power spectrum calculating unit are expanded by said specific vectors ; and

12. A voice recognition method for recognizing an input voice of a narrow frequency band, said voice recognition method including the steps of:

converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band.

13. The voice recognition method as claimed in claim 11, further comprising the steps of  
calculating power spectrums of said input voice of said narrow frequency band;  
calculating expansion coefficients that said power spectrums are expanded by specific vectors ; and  
calculating additional power spectrums in a lack frequency band by use of said expansion coefficients, where said lack frequency band is covered by said wide frequency band but not covered by said narrow frequency band, so as to combine said additional power spectrums in said lack frequency band into said power spectrum of said narrow frequency band, thereby to prepare pseudo power spectrums of said pseudo voice of said wide frequency band.

14. The voice recognition method as claimed in claim 13, further comprising the step of:  
calculating a mercepstrum characteristic quantity based on said pseudo power spectrum.

15. The voice recognition method as claimed in claim 14, further including the steps of:  
comparing said mercepstrum characteristic quantity with a standard pattern for carrying out the voice recognition.

16. A standard pattern preparation method for preparing a standard pattern to recognize an input voice

of a narrow frequency band, said standard pattern preparation method including the steps of:

converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band.

5

**17.** The standard pattern preparation method as claimed in claim 16, further including the steps of:

calculating power spectrums of said input voice of said narrow frequency band;

15

calculating expansion coefficients that said power spectrums are expanded by specific vectors ; and

calculating additional power spectrums in a lack frequency band by use of said expansion coefficients, where said lack frequency band is covered by said wide frequency band but not covered by said narrow frequency band, and combining said additional power spectrums in

20

said lack frequency band into said power spectrum of said narrow frequency band to prepare pseudo power spectrums of said pseudo voice

25

of said wide frequency band.

**18.** The standard pattern preparation method as claimed in claim 17, further comprising the step of:

30

calculating a mercepstrum characteristic quantity based on said pseudo power spectrum.

**19.** The standard pattern preparation method as claimed in claim 18, further including the step of:

35

comparing said mercepstrum characteristic quantity with a standard pattern for carrying out the voice recognition.

40

**20.** A characteristic extraction method for extracting a mercepstrum characteristic quantity of an input voice of a narrow frequency band, said characteristic extraction method including the steps of:

45

calculating power spectrums of said input voice of said narrow frequency band;

converting said input voice of said narrow frequency band into a pseudo voice of a wide frequency band which covers an entirety of said narrow frequency band and which is wider than said narrow frequency band; and

50

calculating a mercepstrum characteristic quantity based on said pseudo voice of said wide frequency band.

55

**21.** The characteristic extraction method as claimed in

claim 20, further comprising the steps of:

calculating expansion coefficients that said power spectrums are expanded by specific vectors ; and

calculating additional power spectrums in a lack frequency band by use of said expansion coefficients, where said lack frequency band is covered by said wide frequency band but not covered by said narrow frequency band, and combining said additional power spectrums in said lack frequency band into said power spectrum of said narrow frequency band to prepare pseudo power spectrums of said pseudo voice of said wide frequency band.

FIG. 1 prior art

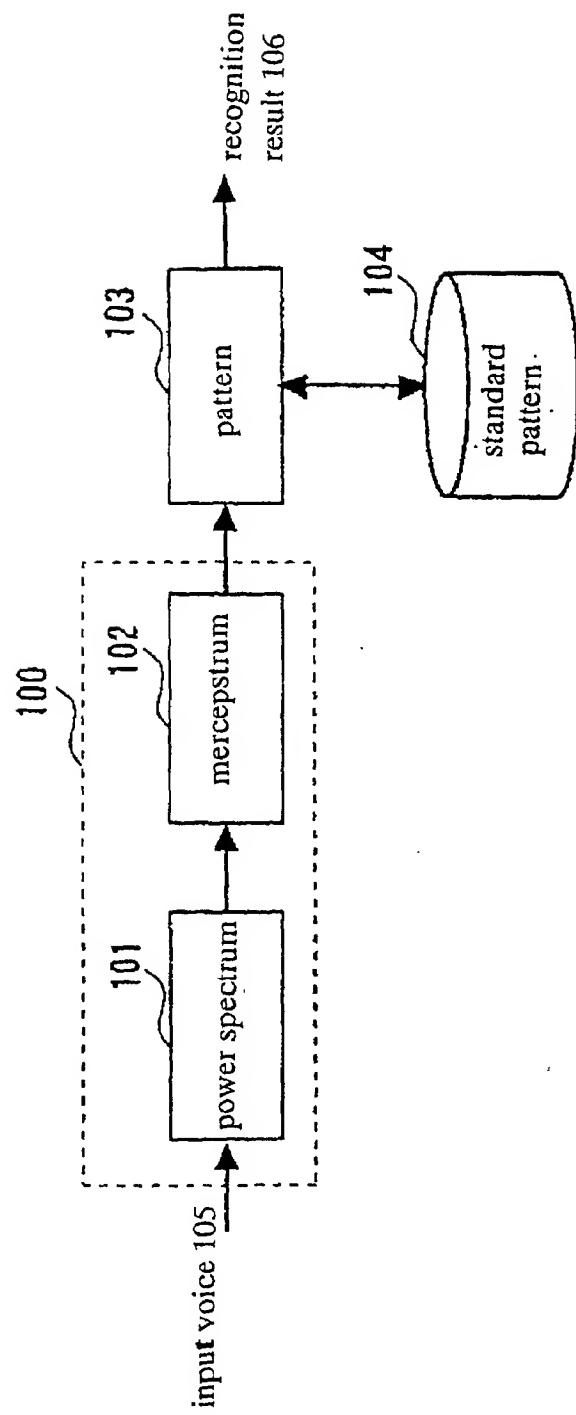


FIG. 2 prior art

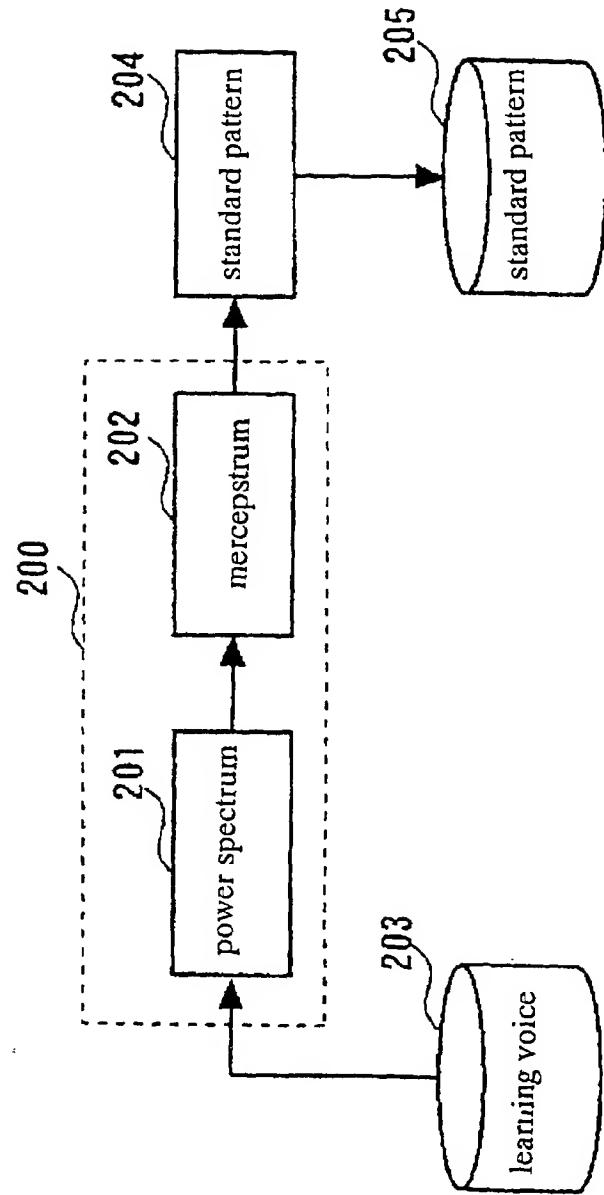


FIG. 3

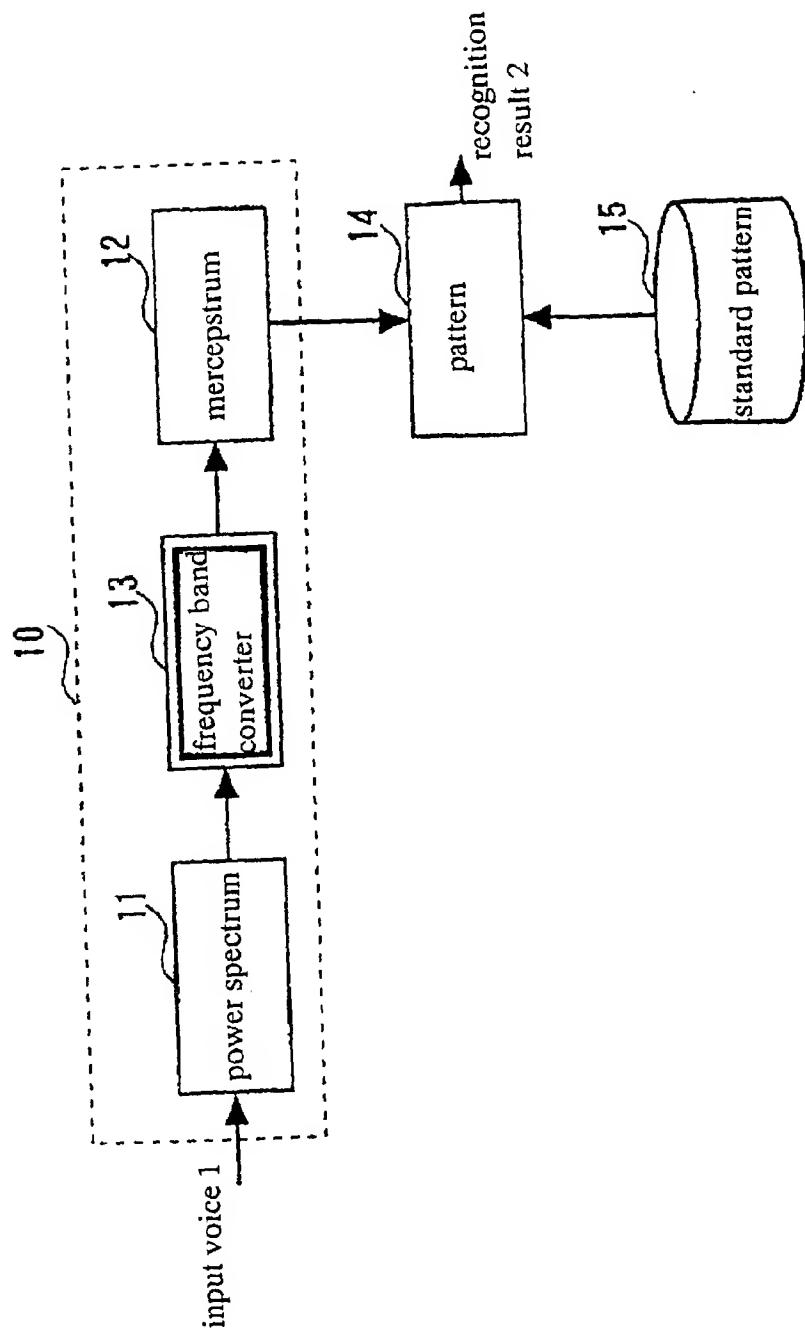


FIG. 4

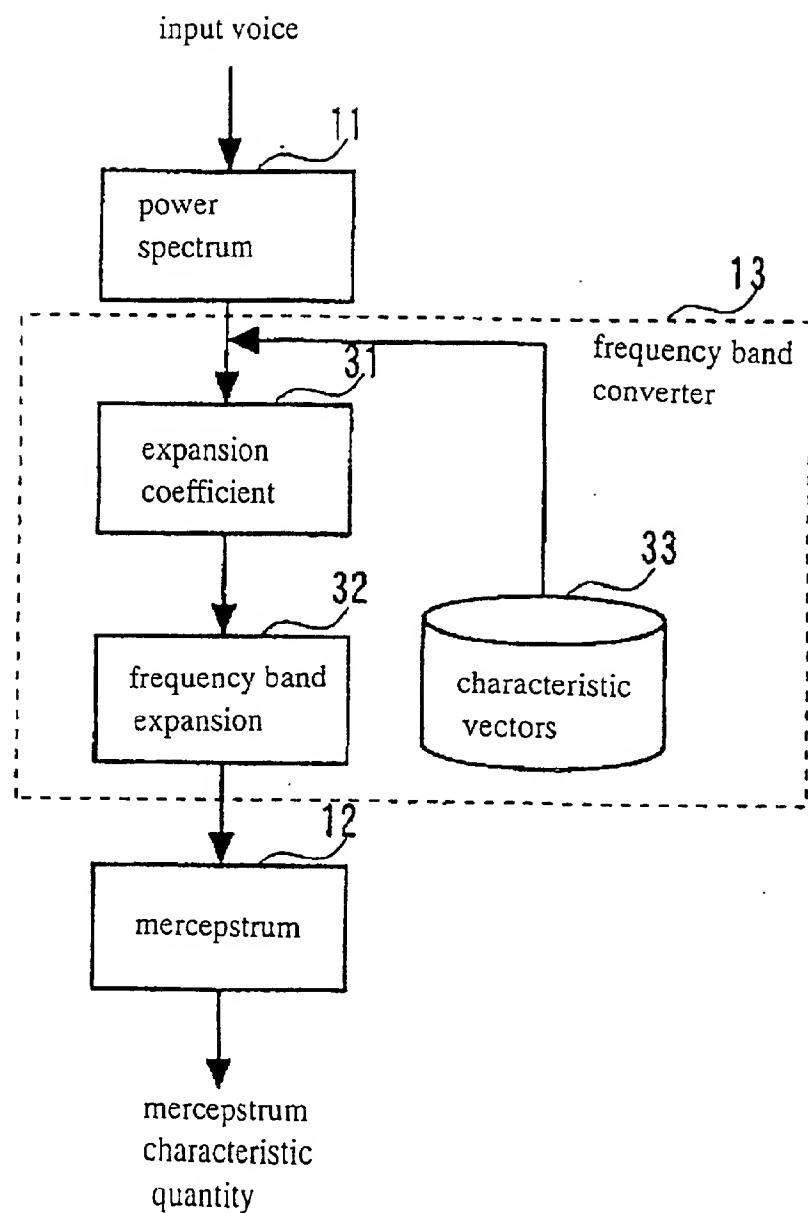


FIG. 5

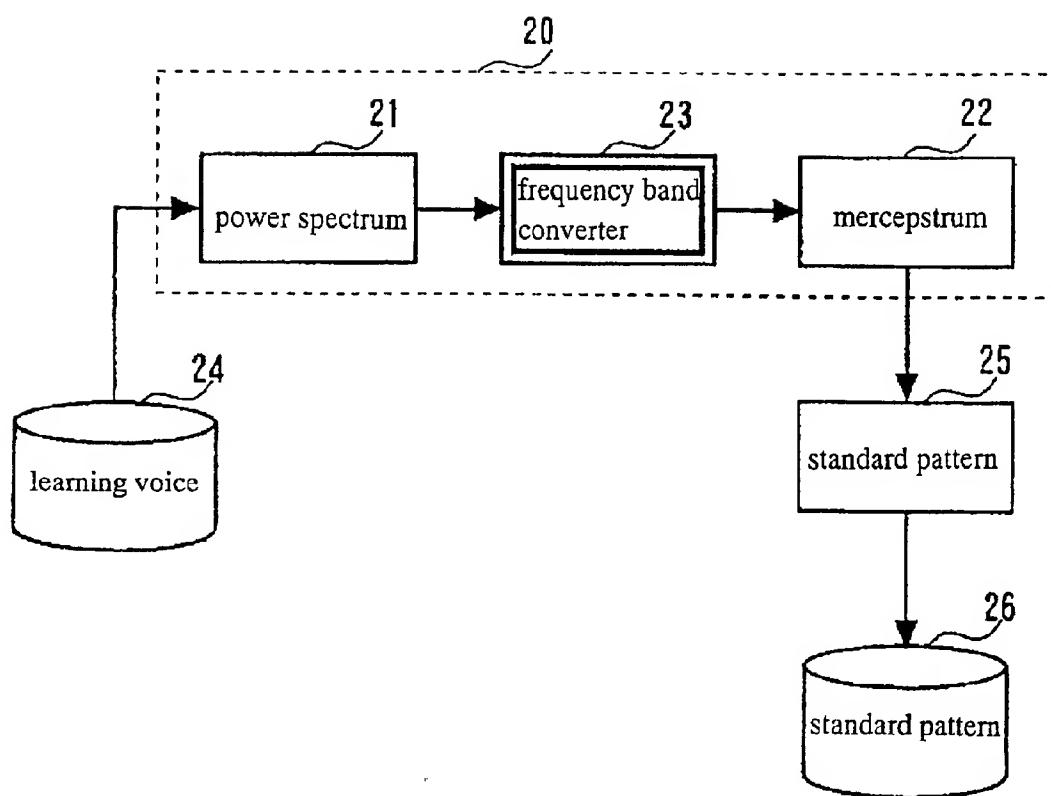


FIG. 6

